The Application of Machine Learning Algorithms in Understanding the Effect of Core/Shell Technique on Improving Powder Compactability

By

Hao Lou

Submitted to the graduate degree program in Pharmaceutical Chemistry and the Graduate Faculty of the University of Kansas in partial fulfillment of the requirements for the degree of Master of Science.

Chair: Michael Hageman

John Chung

Cory Berkland

Date Defended: 11 July 2019

The thesis committee for Hao Lou certifies that this is the approved version of the following thesis:

The Application of Machine Learning Algorithms in Understanding the Effect of Core/Shell Technique on Improving Powder Compactability

Chair: Michael Hageman

John Chung

Cory Berkland

Date Approved: 11 July 2019

Abstract

The study in this thesis systemically investigated the application of core/shell technique to improve powder compactability. A 28-run Design-of-Experiment (DoE) was conducted to evaluate the effects of the type of core and shell materials and their concentrations on tensile strength and brittleness index. Six machine learning algorithms were used to model the relationships of product profile outputs and raw material attribute inputs: response surface methodology (RSM), support vector machine (SVM), and four different types of artificial neural networks (ANN), namely, Backpropagation Neural Network (BPNN), Genetic Algorithm Based BPNN (GA-BPNN), Mind Evolutionary Algorithm Based BPNN (MEA-BPNN), and Extreme Learning Machine (ELM). Their predictive and generalization performance were compared with the training dataset as well as an external dataset. The results indicated that the core/shell technique significantly improved powder compactability over the physical mixture. All machine learning algorithms being evaluated provided acceptable predictability and capability of generalization; furthermore, the ANN algorithms were shown to be more capable of handling convoluted and non-linear patterns of dataset (i.e. the DoE dataset in this study). Using these models, the relationship of product profile outputs and raw material attribute inputs were disclosed and visualized.

Acknowledgments

I would like to express my deepest gratitude to my advisor, Dr. Michael Hageman, for his excellent guidance and training. I am also thankful for the help and encouragement from my colleagues Dr. John Chung, Dr. Yuan-Hon Kiang, and Dr. Ling-Yun Xiao. I would also like to thank Dr. Cory Berkland to be in my thesis committee. Their valuable suggestions enabled me to accomplish this thesis. I would also like to extend my special thanks to my family for their support.

Table of Contents

Chapter 1: Introduction

1.1.Powder Compactability Improvement: Core/Shell Technique vs. Other Techniques

Tablets are the most preferred dosage forms for oral route of administration. However, many active pharmaceutical ingredients (API) possess poor compaction properties, making them difficult to compress into strong tablets. One way to alleviate this issue is to add a compressible excipient such as microcrystalline cellulose (MCC) to the tablet formulation. However, large amounts of excipients in the formulation limit drug loading, which may present a challenge for formulation scientists, particularly if a high drug loading is required to make a reasonably sized, high dose strength tablet for patients. Alternatively, at the molecular level, powder compactability may be improved by modifying the solid state properties of the API. Approaches reported in the literature include forming polymorphs, co-crystals, amorphates, salts, (an)hydrates, etc [1,2,3,4]. Unfortunately, these approaches may not be feasible due to (1) the inability to identify a new form with acceptable compactability; (2) physical instability and sensitivity to moisture, heat, and force during processing conditions, (3) it is highly dependent on the physicochemical properties of the drug, making its application relatively narrow. Powder surface coating (core/shell structure) has been reported to be another strategy with wider potential applicability to improve the compactability of powders [5,6,7,8]. In theory, tablet compaction can be equated to a process of bonding individual particles to a coherent compact under pressure. From a microscopic view, the formation of bonding relies on three main mechanisms: distance forces (i.e., van der Waals forces, hydrogen bonding, electrostatic forces), solid bridges, and mechanical interlocking [9]. A coating layer (shell) may change the powder surface as well as bonding mechanisms. Shi et al. coated polyvinylpyrrolidone (PVP) on fine silicone dioxide particles to convert the compaction behavior from elastic to plastic and

consequently prevent tablet lamination and capping [5]. Furthermore, they coated hydroxypropyl cellulose (HPC) on acetaminophen to improve tabletability [6]. Moreover, by coating polymers on pellets, Yeboah et al. enabled the compression of tablets containing a multiple unit pellet system (MUPS) at low compaction pressures so that the functional coating layer of the pellets was kept intact during the tablet compaction process [8].

1.2.Machine Learning and Its Potential Application in Powder Compactability Study

Despite multiple advantages of the core/shell technique, to our knowledge, the current literature is mainly focused on demonstrating the feasibility of powder compactability enhancement. However, we are not aware of published accounts of systemic evaluations for this technique in terms of its critical material attributes (i.e., type/amount of core/shell material) as well as manufacturability/scalability. Applying a "Quality-by-Design" (QbD) concept can contribute to a systemic, science-driven, and risk-based manner [10] to better understand the core/shell technique and its application in product development. One of the major benefits of "Design-of-Experiment" (DoE) is that it uses less experiments to explore the joint influence of multidimensional inputs (e.g., material attributes, process parameters) on outputs (e.g., product attributes) as compared to univariate experimental design.

Additionally, predictive models can be developed in attempt to accurately describe the relationship of inputs and outputs in the design space. Response surface methodology (RSM) is currently the most widely-employed model for formulation/process development and listed in ICH Q8 guidance [10]. RSM models can (1) be derived by fitting experimental data into a polynomial function containing quadratic terms and (2) provide predicted values using the obtained equations. With improvements in computing speed and new concepts proposed in the field of machine learning, additional algorithms have been developed to handle larger and more

complicated datasets and provide more accurate predictions. One example is Support Vector Machine (SVM), an algorithm based on transforming data into a hyperplane and is currently being applied in drug design and molecular assessment [11,12,13]. Another example is Artificial Neural Network (ANN), a digitized model emulating the human brain's neurological behavior, which displays superhuman performance in some fields [14]. Specifically, in the pharmaceutical field, ANN has been attempted to be used in drug design, molecular modeling, formulation development, quality assessment, and process validation [15,16,17,18,19]. In contrast to traditional statistical models such as RSM, ANN is more powerful in terms of (1) handling large, convoluted, multi-dimensional, and non-linear patterns of databases; (2) performing multi-objective optimization; (3) continuously improving predictability with more training datasets [18.19]. Many types of ANN using various topologies and algorithms have been developed but their applications in pharmaceutical research have not been widely explored.

1.3.Specific Objective of This Thesis

In this study, first, a scalable powder coating (core/shell) manufacture route (spray drying) was developed to improve powder compactability. Two brittle excipients, Foremost™ Lactose 313 and Emcompress® Calcium Phosphate Premium powder [20], were selected as surrogates of many brittle APIs. Next, critical material parameters (CMP) were thoroughly investigated by a 4-factor 28-run DoE study. Furthermore, based on the DoE data as the training dataset, predictive models were developed by using various established algorithms, including Response Surface Methodology (RSM), Support Vector Regression (SVR), and four types of Artificial Neural Network (ANN) algorithms: Backpropagation Neural Network (BPNN), Genetic Algorithm Based BPNN (GA-BPNN), Mind Evolutionary Algorithm Based BPNN (MEA-BPNN), and Extreme Learning Machine (ELM). All these obtained models helped to

understand the effect of critical material parameters of core/shell technique on powder compactability and explore the utility of such algorithms in a pharmaceutical formulation and processing environment.

Chapter 2: Materials and Methods

2.1.Materials

Lactose 313 was purchased from Foremost Farms USA (WI, USA). Kollidon® K30 (PVP) and Kollidon® VA64 (VA64) were kindly donated by BASF (NY, USA). Emcompress® Premium Powder (Dibasic Calcium Phosphate) was kindly donated by JRS Pharma (NY, USA). Anhydrous ethanol was purchased from Fisher Scientific (MZ, USA).

2.2.Preparation of Physical Mixtures (PM)

To prepare lactose/PVP and lactose/VA64 physical mixtures (PM), all powders were first manually screened through 30-mesh sieves. Subsequently preset amount of sieved powders was precisely weighed and then blended for 5 minutes using a Turbula mixer (Glenn Mills Inc, NJ, USA).

2.3.Preparation of Core/Shell powders

Core/Shell powders were prepared via spray drying. Briefly, according to quantity listed on Tables 1 and 2, preset amount of PVP or VA64 was solubilized in ethanol. Subsequently, lactose or calcium phosphate were added to form a suspension, followed by continued stirring with a magnetic stir bar for at least one hour. The obtained suspension was then spray dried using a mini-spray dryer B-290 (Buchi, Flawil, Switzerland). Spray drying parameters were set as following: drying nitrogen gas 0.55 kg/min, aspirator 90%, inlet temperature 95°C, nozzle gas 7.5 slpm, nozzle cooling temperature 20°C. Typical batch sizes were from 12-25 grams sprayed over 45-90 minutes.

2.4.Particle Size Measurement

Particle size was characterized using a HELOS laser diffractor (Sympatec, Clausthal-Zellerfeld, Germany) under RODOS models (dry dispersion state). Particle size distribution (PSD) was numerically expressed by D10, D50, D90, and volume median diameter (VMD).

2.5.Tablet Compression

The harvested powders (PM or Core/Shell powders) were compressed into tablets using a Carver hydraulic press (Carver Inc, IN, USA) with a split die and round 10-mm diameter flat face tooling. Prior to compression, the die and punch were treated with external lubrication using a suspension of magnesium stearate in ethanol (5% w/v). Tablet weight was controlled within a range of 500±20 mg. The compaction force was set to 3000 LB, which was equivalent to 170 MPa of compaction pressure. Tablets were relaxed in sealed bags at room temperature for at least 24 h prior to further characterization.

2.6.Measurements of Tablet Tensile Strength and Tablet Brittleness Index

Tablet tensile strength and brittleness index were measured by a Texture analyzer (Texture Technologies Corp., NY, USA). Only tablets with tensile failure breaking from the central vertical line were chosen for measurements. Testing speed was set at 0.02 mm/s and data collection was initiated by trigger force at 20 g. TS and BI were calculated according to Eq.1 and Eq.2 [21], respectively.

$$TS = \frac{2 \times F}{\pi \times D \times H} \quad \text{(Eq.1)}$$

$$BI = \frac{D}{MEDL} \quad \text{(Eq.2)}$$

Where F is the breaking force, D is the tablet diameter, H is the tablet thickness, and MEDL is the maximum elastic deformation length. To obtain MEDL, a first derivative method [21] was employed to determine the linear segment of the force-displacement curve.

2.7.Morphology

Samples for SEM analysis were prepared by laying a monolayer of sample on a Carbon double sided adhesive tape attached to an Aluminum stub. The sample was sputter-coated with Gold/Palladium (Au/Pd) target using the Cressington 108auto/SE (Cressington Scientific Instruments, Watford, UK). Electron micrographs were taken using the Zeiss EVO-15 system (Carl Zeiss, Oberkochen, Germany) with a LaB6 source.

2.8.Experimental Design

2.8.1. Preliminary Experimental Design

Preliminary studies were performed to compare the compactability using a core/shell strategy vs. binary physical mixture strategy. Compactability was evaluated at different wt/wt% of (1) physical mixtures consisted of lactose and PVP (or VA64); and (2) core/shell powders consisted of lactose (core material) and PVP (shell material).

2.8.2. Design of Experiment (DoE)

Categorical factors (shell material type and core material type) and continuous factors (shell material concentration and core material concentration) were evaluated with their influence on tablet tensile strength and brittleness index (responses) throughout a 28-run DoE study, as shown in Table 1. PVP and VA64 were selected as two types of shell materials. Lactose and Calcium Phosphate were selected as two types of core materials. Out of 28 runs, the first 24 runs contained two categorical factors (2 levels) and two continuous factors (3 levels) designed using Custom Design following the D-optimality criterion by JMP (SAS Institute Inc., NC, USA). Runs 25-28 were added to cover four more axial points, where Runs 25-26 were already performed in the preliminary study and Runs 27-28 contained the levels complimentary to Runs

25-26. For each run, tensile strength and brittleness index values were measured for at least five tablets and all the data were inputted for establishing the model as described in Section 2.8.

Table 1

Design-of-Experiment for four factors including two continuous factors and two categorical factors.

DoE Run #	Conc. of Material for Shell (% wt/v)	Conc. of Material for Core (% wt/v)	Type of Material for Shell	Type of Material for Core
1	0.1	10	S1	C1
2	0.1	10	S2	C2
3	0.1	15	S1	C2
4	0.1	15	S2	C1
5	0.1	20	S1	C1
6	0.1	20	S1	C2
7	0.1	20	S2	C1
8	0.1	20	S2	C2
9	1.05	10	S1	C2
10	1.05	10	S2	C1
11	1.05	15	S1	C1
12	1.05	15	S1	C2
13	1.05	15	S2	C1
14	1.05	15	S2	C2
15	1.05	20	S1	C1
16	1.05	20	S2	C2
17	2.0	10	S1	C1
18	2.0	10	S1	C2
19	2.0	10	S2	C1
20	2.0	10	S2	C2
21	2.0	15	S1	C1
22	2.0	15	S2	C2
23	2.0	20	S1	C2
24	2.0	20	S2	C1
25	0.5	20	S1	C2
26	0.5	20	S2	C2
27	1.5	10	S1	C1
28	1.5	10	S2	C1

S1: Kollidon®30

S2: Kollidon®VA64

C1: Foremost™ Lactose 313

C2: Emcompress® Calcium Phosphate Premium Powder

2.8.3. Model Validation Points

Four validation points within the design space, as shown in Table 2, were selected to validate the robustness of the obtained models from DoE studies and adjust the hyperparameters if needed. For each run, tensile strength and brittleness index values were measured for at least five tablets.

Table 2

Dataset for model validation.

Validation Run #	Conc. of Material for Shell (% wt/v)	Conc. of Material for Core (% wt/v)	Type of Material for Shell	Type of Material for Core
1	0.5	17.5	S1	C2
2	1.5	12.5	S2	C2
3	0.5	12.5	S2	C1
4	1.5	17.5	S1	C1

S1: Kollidon®30

S2: Kollidon®VA64

C1: Foremost™ Lactose 313

C2: Emcompress® Calcium Phosphate Premium Powder

2.9.Predictive Modeling and Learning Algorithms

2.9.1. Response Surface Methodology

Response surface methodology (RSM) generally employs a second-order polynomial regression equation to describe the relationship between factors and responses. In this study, eight quadratic equations were developed, in which every combination of one core material (lactose or Calcium Phosphate), one type of shell material (PVP or VA64), and one type of response (tensile strength or brittleness index) correlated to one quadratic equation. One quadratic equation contains two independent variables: shell material concentration and core material concentration for each type of response, and takes the form of Eq.3:

$$Y = k + aX_1 + bX_2 + cX_1X_2 + dX_1^2 + eX_2^2 \quad \text{(Eq.3)}$$

where Y is the response, X_1 is the concentration of shell material, and X_2 is the concentration of core material.

2.9.2. Support Vector Regression

The algorithm for support vector regression (SVR) intends to find a function that predicted a response value that deviated from their corresponding observed value by a range up to ε and simultaneously this function needs to be as flat as possible [22]. In this study, nonlinear SVR algorithm was utilized to find the flattest function in a feature space and the function could then be restated as to minimize L(α) shown in Eq.4.

$$L(\alpha) = \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}(\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*)G(x_i, x_j) + \varepsilon\sum_{i=1}^{N}(\alpha_i + \alpha_i^*) - \sum_{i=1}^{N} y_i(\alpha_i - \alpha_i^*) \quad \text{(Eq.4)}$$

subject to the constraints $\sum_{i=1}^{N}(\alpha_i - \alpha_i^*) = 0$ and α_i , α_i^* within [0, C]

where L(α) is the Lagrange dual function, α_i , α_i^* , α_j , α_j^* are Lagrange multipliers, C is a constant determining the trade-off between the function flatness and the tolerance of largest deviation to ε, $G(x_i, x_j)$ is the Kernel function to map the training dataset from an input space to a high-

dimensional feature space. Amongst different Kernel functions, the Gaussian radial base function (RBF) (Eq.5) was selected for our SVR model.

$$G(x_i, x_j) = \exp(-\frac{||x_i - x_j||^2}{2\sigma^2}) \quad \text{(Eq.5)}$$

where x_i and x_j are two feature vectors, and $\frac{1}{2\sigma^2}$ (Y) is the inverse of the standard deviation.

Parameters C and Y were optimized through a grid search algorithm and set at 50000 and 0.006, respectively.

2.9.3. Backpropagation Neural Network

The backpropagation algorithm is a type of multilayer feedforward neural network. In this study, two three-layer BPNNs were developed for tensile strength and brittleness index respectively using a training dataset containing 165 samples and the structure was shown in Figure 1. The outputs of hidden layer and output layer were computed according to Eq. 6.

$$a_j^l = \sigma(\sum_k w_{jk}^l a_k^{l-1} + b_j^l) \quad \text{(Eq.6)}$$

where w_{jk}^l is the weight interconnecting the k^{th} node in $(l-1)^{th}$ layer and the j^{th} node in l^{th} layer, b_j^l is the threshold of the j^{th} node at l layer, a_k^{l-1} is the output of the k^{th} node in $(l-1)^{th}$ layer, a_j^l is the output of the j^{th} node in l^{th} layer, and σ is the transfer function.

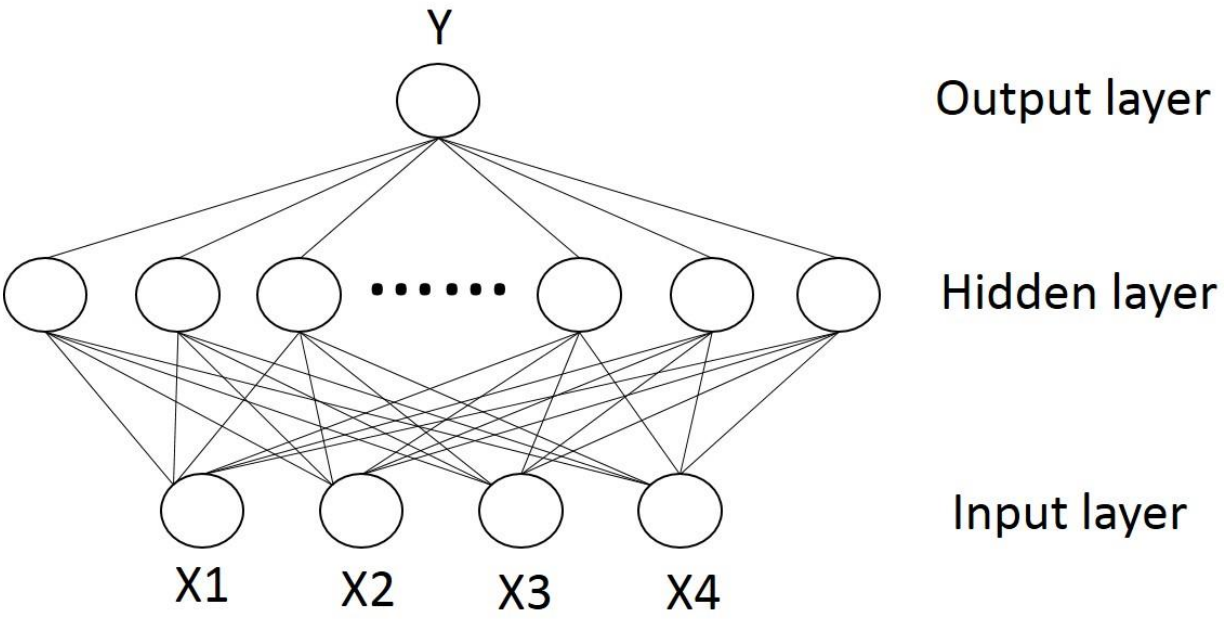

Fig.1. Structure of a three-layer neural network.

To determine the desired number of nodes at hidden layer, Eq.7 was employed:

$$N_{hidden}=\frac{N_{sample}/\beta - N_{output}}{N_{input}+N_{output}+1} \quad \text{(Eq.7)}$$

since β (the degree of determination of the network) was recommended to be greater than 1 for the prevention of overfitting [19], maximal N_{hidden} was set to 27 and a screening was conducted from 10 to 27. Eventually 19 nodes were selected for the hidden layer. In addition, a sigmoid function was selected for the transfer function between input layer and hidden layer and a linear function was selected for the transfer function between the hidden layer and output layer. The performance of BPNN was measured by an error function, which is the mean squared error between predicted values and observed values. Throughout N iterative learning steps (N=50000 for our programming), all weight and threshold values were adjusted after each step to make predicted values proximal to actual values. To prevent overfitting, a cross-validation was carried out by dropping out 10% of the data from the training set for each loop.

2.9.4. Genetic Algorithm Based Backpropagation Neural Network

In general, for BPNN algorithm, the weights and thresholds were initialized randomly. Whereas, using GA-BPNN for this study, the initial weights and thresholds were given through a selection process based on Genetic Algorithm (GA), whose flow was represented in Figure 2. For our programming, all weights and thresholds were encoded to a chromosome which contained 115 strings and 50 chromosomes were randomly generated as the initial population. According to the smallest adaptive value (i.e., the mean square error between predicted value and the actual value) as the criteria, the best chromosome of one generation was selected. Next, a new population of 50 chromosomes were generated by inheriting the well-adapted chromosomes from the old population, followed by certain probability of crossover (0.2) and mutation (0.05) operations at some random points in the chromosomes. Again, a new best adaptive value would be found from this generation and the old one would be replaced if this new one was more adaptive. Overall ten generations were iterated in our programming. Eventually, the optimized chromosomes were selected and subsequently utilized for the initialization of BPNN training. Noteworthy, regarding BPNN topology, log-sigmoid function was selected as the transfer function between input layer and hidden layer and other neural network structural parameters were selected the same as those of BPNN.

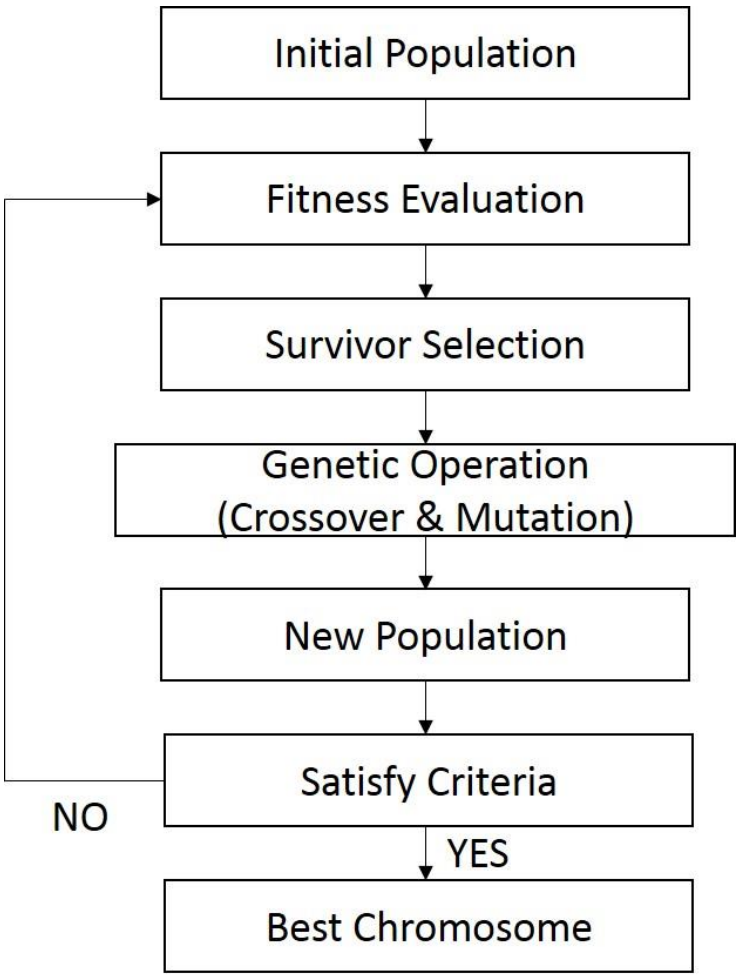

Fig.2. The flow diagram of Genetic Algorithm.

2.9.5. Mind Evolutionary Algorithm Based Backpropagation Neural Network

Mind evolutionary algorithm based backpropagation neural network (MEA-BPNN) was developed with the same purpose as GA-BPNN in order to give initialized weights and thresholds. In terms of algorithm, MEA-BPNN simulated human being's mind evolution and utilized similar-taxis and dissimilation operations [23,24] to substitute crossover and mutation of GA-BPNN. A flow of MEA-based algorithm was represented in Figure 3. One individual, similar as chromosome of GA-BPNN, contained the information of one set of weights and thresholds. For our algorithm, briefly, 5 superior groups and 5 temporary groups were initialized

randomly with 20 individuals in each group. Next, local competition was carried out through similar-taxis operation and global competition was carried out through dissimilation operation. With the continual generation of new groups and discard of old groups, the above-mentioned operations were iterated for 10 times. In the end, the winner superior group with the highest score (i.e., the inverse of the mean square error between predicted value and the actual value) was updated to global billboard as the global optimum. The optimized individual was utilized for the initialization of BPNN training. Noteworthy, regarding BPNN topology, log-sigmoid function was selected as the transfer function between input layer and hidden layer. Additionally, 14 nodes were used for hidden layer for tensile strength modeling and 16 nodes were used for hidden layer for brittleness index modeling. Other neural network structural parameters were selected the same as those of BPNN.

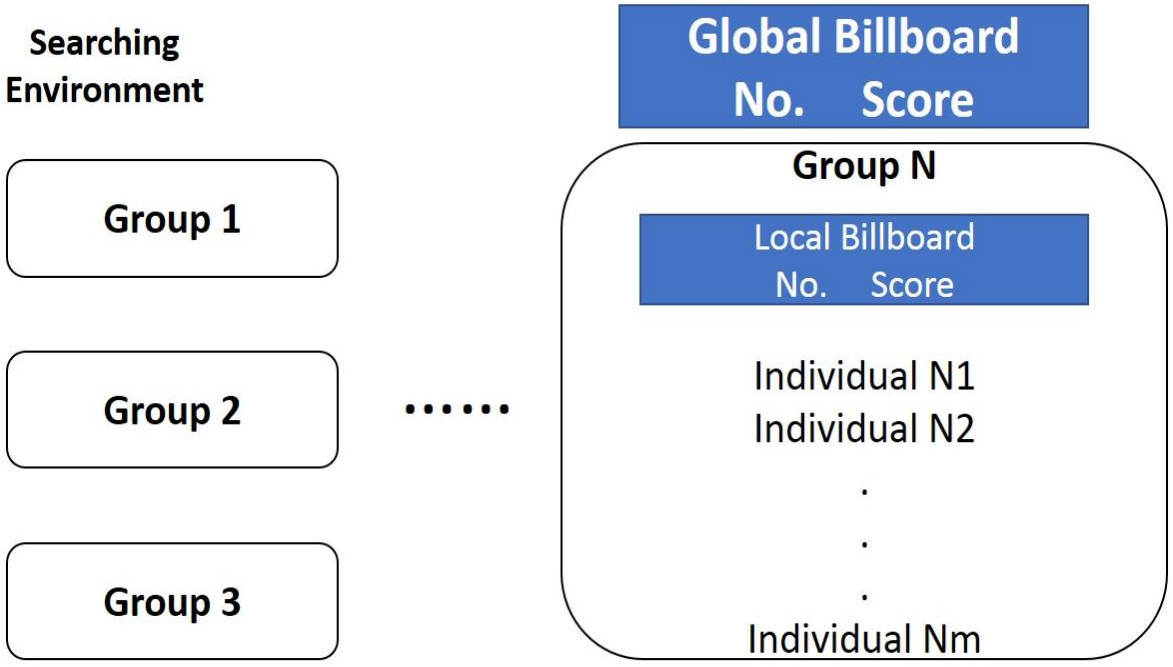

Fig.3. The flow diagram of Mind Evolutionary Algorithm.

2.9.6. Extreme Learning Machine

In this study, extreme learning machine (ELM) was built with the similar structure as BPNN (i.e., three layers, four features in input layer, one feature in output layer, sigmoid function as the transfer function between input and hidden layers, as well as linear function as the transfer function between hidden and output layers). However, unlike iterative steps of adjusting weights and thresholds for BPNN learning, the algorithm of ELM assigned hidden nodes randomly and only needed to calculate the hidden layer output matrix H with one learning step [25]. This difference in algorithm lead to more nodes in hidden layer for ELM. As deciding the node number $\tilde{N}$ ($\tilde{N} \leq$ sample number N) in hidden layer, the learning target for zero error was followed [25], with $\sum_{j=1}^{\tilde{N}} \| o_j - t_j \| = 0$ (o_j: observed output; t_j: predicted output), which can also be written as $H\beta = T$ (H: hidden layer output matrix; β: weight vector; T: predicted output vector) [25]. For our algorithm, $\tilde{N}$ was optimized from 15 to 50 and eventually 31 was selected.

2.9.7. Statistical Analysis

To evaluate the difference between observed values and predicted values, coefficient of variance of the root mean square error "CV(RMSE)" were applied.

In addition, Akaike information criterion "AIC" was employed to evaluate the optimality and relative quality of all neural network models, according to Eq.8.

$$AIC = n_s \times \ln(SS) + 2 \times n_w \quad \text{(Eq.8)}$$

where n_s is the number of datasets, n_w is the number of weights and nodes, and SS is the residual sum of squares between observed and predicted values.

Chapter 3: Results and Discussion

3.1.Core/Shell Powder vs. Physical Mixture

As the compactability of core/shell powder was compared to the compactability of physical mixture, the tensile strength of the lactose/polymer compacts was plotted against the weight percentage of the compactible polymer (Figure 4). For the physical mixture, increasing PVP did not show any improvement of the tensile strength until it reached 50% (wt/wt). The maximal tensile strength (5.05 MPa) was observed at around 90% (wt/wt) of PVP in the physical mixture. Mixing VA64 with lactose reduced the amount of polymer required to initiate the compactibility improvement to 17% (wt/wt) of polymer in the physical mixture. Maximum tensile strength for the lactose/VA64 physical mixture was again observed at 90% (wt/wt) of polymer, with a tensile strength of 7.26 MPa achieved, almost doubling the increase provided by that of the PVP, indicating VA64 powders improved compactability of physical mixture more significantly.

In Figure 4, the core/shell technique not only reduced the quantity of the PVP needed to initiate the tensile strength increase, but also improved tensile strength at a much greater rate compared to the physical mixture. As seen, a tensile strength of 3.52 MPa was achieved with a PVP weight percentage of 2.5% (wt/wt) using the core/shell technique. By contrast, approximately 60% (wt/wt) of PVP was required to achieve the same level of tensile strength physically mixing PVP with lactose. To further increase the tensile strength from 3.52 MPa to 4.33 MPa, only 6.5% (wt/wt) additional PVP (total 9% PVP) was needed using the core/shell technique whereas approximately 10% (wt/wt) additional PVP (total ~70% PVP) was needed for the physical mixture.

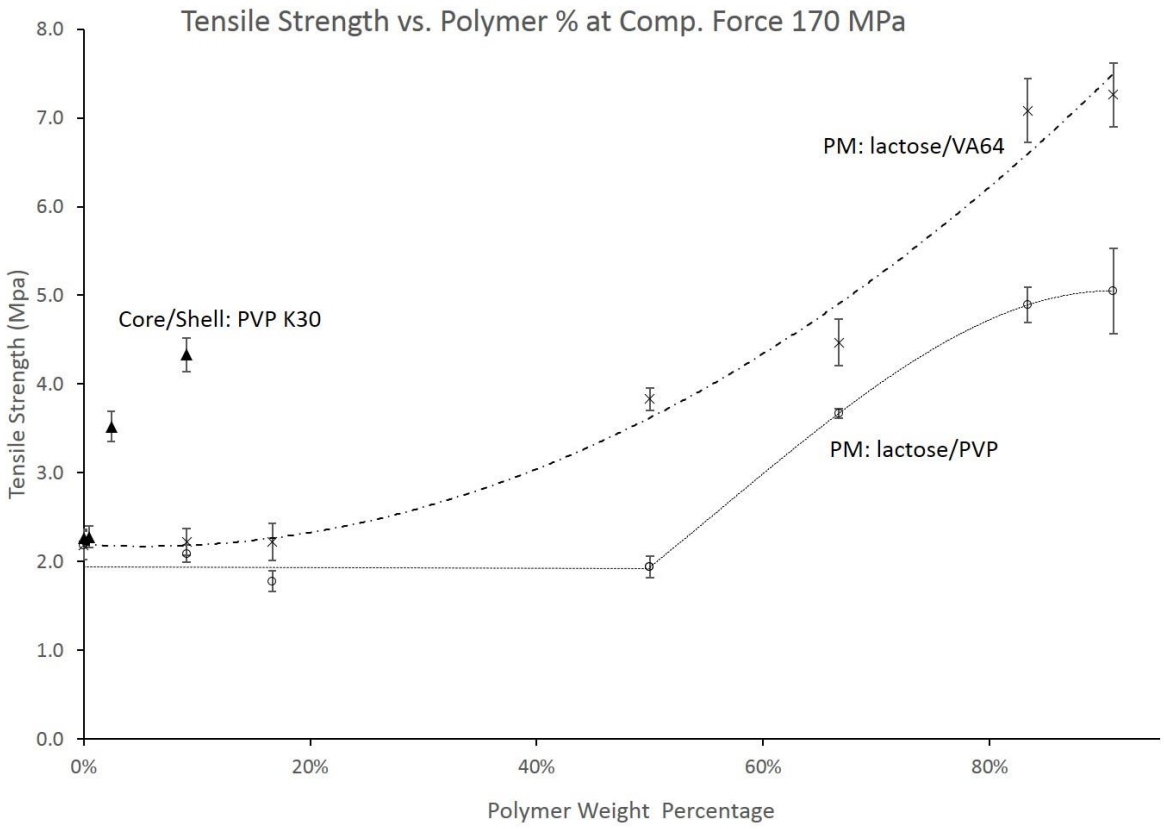

Fig.4. Effect of polymer weight percentage on tensile strength.

Symbol °: physical mixture of lactose/PVP

Symbol ˣ: physical mixture of lactose/VA64

Symbol ▲: core/shell of lactose/PVP

In addition to the tensile strength at a given compaction pressure, brittleness index of a material is another indicator of its compactibility, as it is generally accepted that less brittle materials are more compactible. Figure 5 show a general trend of decreasing brittleness index with increasing amount of PVP. The core/shell technique significantly reduced the amount of PVP required compared to the physical mixture. Both the tensile strength and brittleness index results demonstrated that the core/shell technique is a better approach to improve compactibility

of a poorly compactible API for stronger and less defect prone tablets while maximizing potential drug load in a tablet.

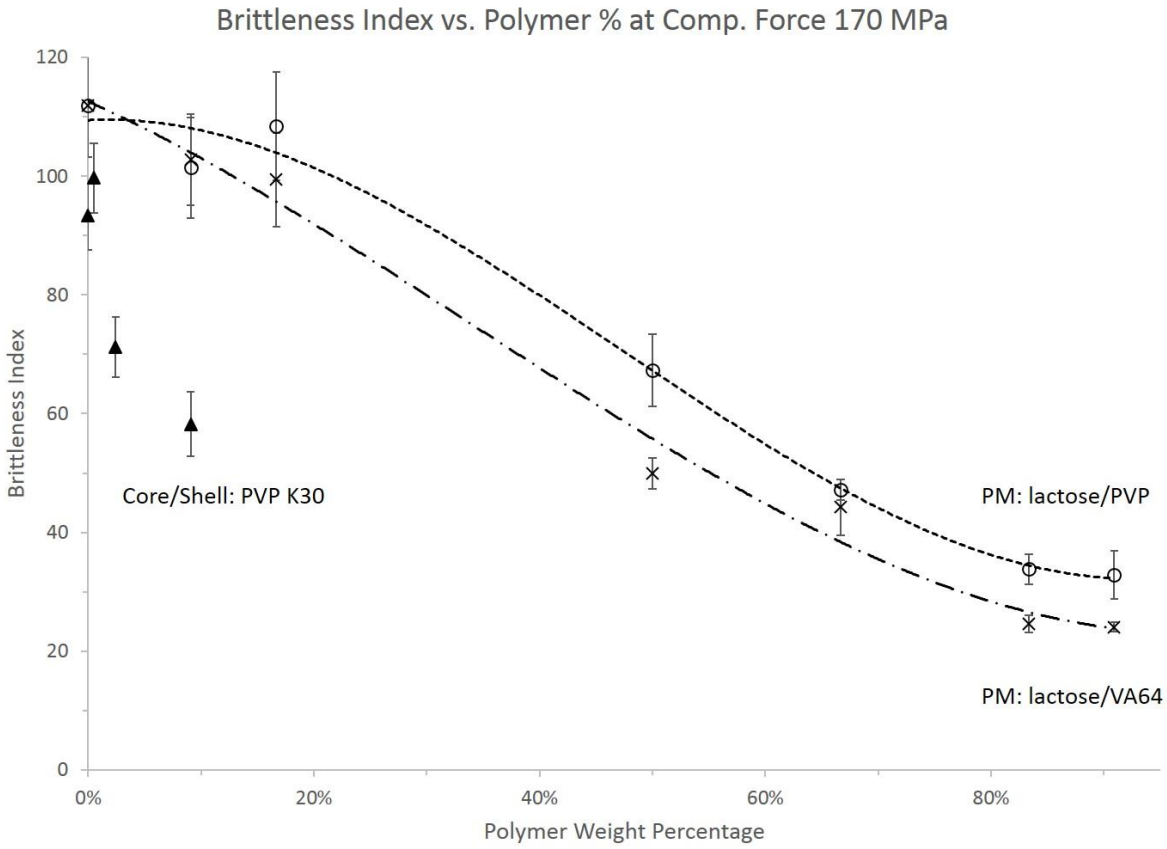

Fig.5. (a) Effect of polymer weight percentage on brittleness index.

Symbol °: physical mixture of lactose/PVP

Symbol ˣ: physical mixture of lactose/VA64

Symbol ▲: core/shell of lactose/PVP

In this study, core/shell powders were prepared using a continuous and scalable spray drying process. To evaluate this process, particle size distribution (PSD) was monitored before and after spray drying to assure that PSD of the powders for core was maintained during the spray-drying process, and thus consistent with and comparable to the lactose PSD in the physical

mixtures. Table 3 showed that using the given process parameters, there was no significant PSD change observed after spray drying, indicating that no agglomeration or granulation occurred during the core/shell preparation process. The slight decrease in PSD of the core/shell powders may be hypothesized to be attributed to fines of the free polymer (i.e. PVP), which is not absorbed onto the surface of the lactose particles, generated by spray drying.

Table 3

Particle size distribution of core/shell powders prepared by spray drying.

Material	D10 (μm)	D50 (μm)	D90 (μm)	VMD (μm)
Lactose	9.3	37.6	62.7	38.6
Spray-dried lactose from EtOH	9.2	36.9	60.8	36.3
Spray-dried lactose/PVP (20%/0.1%)	8.7	34.5	52.0	32.5
Spray-dried lactose/PVP (20%/0.5%)	8.2	31.8	56.5	32.8
Spray-dried lactose/VA64 (20%/0.5%)	8.5	33.5	57.5	33.9
Spray-dried lactose/VA64 (20%/2%)	8.0	31.0	53.0	31.1
Spray-dried lactose/VA64 (10%/2%)	7.5	28.4	53.2	29.3

%: wt/v in EtOH

3.2.Design-of-Experiment

The experimental results of the Design-of-Experiment (DoE) study were presented in Tables 5 and 6. Among the factor levels explored, a wide range of tensile strength and brittleness index values were observed, i.e., the lowest tensile strength values were 2.18 MPa (Run 2) and 2.14 MPa (Run 8) whereas the highest tensile strength were 7.58 MPa (Run 21) and 7.63 MPa (Run 27). In addition, powders with a low tensile strength were more brittle and those with high tensile strength were more plastic. In Figure 6, the relationship of tensile strength and brittleness index could approximately be fit to a power law function $BI = C_0 \times TS^{C_1}$, where C_0 and C_1 were constants, BI was brittleness index, and TS was tensile strength. This observation was consistent with the findings from Gong et al [27]. Another interesting result was that when the same weight ratio of dispersed core material to solubilized shell material was used for spray drying, tensile strength and brittleness index values might not be the same and was dependent on the polymer type and concentration. For example, with the same weight ratio at around 10 to 1, regarding lactose/PVP combination, Run 9 (5.30 MPa) and Run 23 (4.33 MPa) were different, whereas regarding calcium phosphate/VA64 combination, Run 10 (4.85 MPa) were similar as Run 24 (4.68 MPa). This observation was speculated to be related to the polymer (shell material) molecular structure and its ability of adsorbing at the solid-liquid interface. From another perspective, core material solubility in shell solution can influence compressibility. For example, Sugimori et al. found that high drug solubility in binder solution lowered compressibility of granules coated with binder [28]. Nevertheless, in this study, the factor of core material solubility is not considered as a profound factor since both lactose and calcium phosphate are insoluble in ethanol. Overall, these results showed the chosen parameters as well as their levels within the DoE design space significantly influenced responses. In addition, this observation provided an

insight that core/shell technique could achieve different purposes depending on the properties of shell materials, e.g., our study demonstrated the improvement of compactability with PVP and VA 64, whereas coating with materials such as Magnesium Stearate or Silica could improve flowability [29, 30].

Table 5

Observed tensile strength values from experiments and predicted tensile strength values from models.

Do E Run #	Tensile Strength (MPa)						
	Actual	Predicted (RSM)	Predicted (SVM)	Predicted (ELM)	Predicted (BPNN)	Predicted (GA-BPNN)	Predicted (EMA-BPNN)
1	2.45±0.31	2.54	2.48	2.45	2.45	2.74	2.41
2	2.18±0.07	2.19	2.18	2.18	2.18	2.18	2.24
3	2.42±0.14	2.32	2.50	2.42	2.42	2.55	2.42
4	2.43±0.22	2.41	2.60	2.43	2.43	2.42	2.50
5	2.51±0.23	2.34	2.66	2.51	2.51	2.60	2.46
6	2.28±0.12	2.48	2.32	2.28	2.32	2.31	2.25
7	2.42±0.19	2.42	2.44	2.42	2.43	2.42	2.34
8	2.14±0.15	2.05	2.08	2.14	2.13	2.10	2.12
9	5.30±0.65	5.35	5.26	5.30	5.30	5.44	5.37
10	4.85±0.57	4.92	5.06	4.85	4.64	4.88	4.97
11	5.67±0.97	6.02	5.90	5.67	5.67	5.72	5.67
12	4.26±0.16	4.36	4.74	4.26	4.26	4.26	4.19
13	5.87±0.53	5.89	5.20	5.87	5.87	5.89	6.03
14	3.63±0.23	3.60	3.77	3.63	3.67	3.63	3.49

15	5.38±0.60	5.55	5.42	5.39	5.39	5.54	5.53
16	3.27±0.12	3.18	3.17	3.27	3.28	3.22	3.23
17	6.83±0.41	7.41	7.49	6.84	6.83	6.87	6.81
18	6.65±0.62	6.60	6.19	6.65	6.44	6.41	6.65
19	6.04±0.46	6.08	6.06	6.04	6.14	6.20	5.98
20	4.02±0.21	4.01	4.21	4.02	3.98	3.92	3.97
21	7.58±0.42	7.23	7.28	7.58	7.58	7.58	7.57
22	4.29±0.11	4.32	4.07	4.29	4.30	4.29	4.31
23	4.33±0.19	4.35	4.41	4.33	4.33	4.36	4.36
24	4.68±0.54	4.68	4.83	4.68	4.68	4.75	4.79
25	3.52±0.17	3.30	3.21	3.52	3.57	3.50	3.52
26	2.39±0.17	2.57	2.68	2.39	2.40	2.42	2.34
27	7.63±0.73	6.96	6.92	7.63	7.63	7.59	8.01
28	5.97±0.74	5.86	5.76	5.97	5.97	5.83	5.81
		CV(RMSE): 4.9%	CV(RMSE): 6.8%	CV(RMSE): 0%	CV(RMSE): 1.4%	CV(RMSE): 2.3%	CV(RMSE): 2.4%
				AIC: 230	AIC: 162	AIC: 174	AIC: 115

RSM: Response Surface Methodology

SVM: Support Vector Machine

BPNN: Backpropagation Neural Network

GenBPNN: Backpropagation Neural Network based on Genetic Algorithm

EvoBPNN: Backpropagation Neural Network based on Evolutionary Algorithm

ELM: Extreme Learning Machine

CV(RMSE): Coefficient of Variance of the Root Mean Squared Error

AIC: Akaike Information Criterion

Table 6

Observed brittleness index values from experiments and predicted brittleness index values from models.

DoE Run #	Brittleness Index						
	Actual	Predicted (RSM)	Predicted (SVM)	Predicted (ELM)	Predicted (BPNN)	Predicted (GA-BPNN)	Predicted (EMA-BPNN)
1	111.6±11.2	89.1	108.1	111.6	111.6	98.01	110.4
2	106.2±5.3	106.1	109.6	106.2	108.5	100.32	110.0
3	95.1±7.4	96.8	94.2	95.1	95.1	96.26	98.1
4	119.2±14.9	118.9	110.1	119.2	119.2	121.80	124.0
5	108.1±11.1	108.5	103.5	108.1	108.1	103.09	102.0
6	99.7±5.9	96.2	93.5	99.7	98.1	98.30	99.2
7	115.5±8.7	115.6	111.9	115.5	116.6	116.44	113.7
8	104.4±6.7	105.7	109.1	104.4	104.4	110.34	104.3
9	47.9±5.9	47.1	47.4	47.9	46.3	52.20	47.3
10	55.3±4.0	56.6	54.6	55.3	56.7	52.17	55.2
11	52.2±4.6	51.4	49.6	52.2	52.2	52.71	52.8
12	55.3±5.4	53.6	51	55.3	55.3	57.51	56.7
13	49.2±4.5	49.5	55.9	49.2	49.2	46.58	51.2
14	66.4±1.9	66.8	64.9	66.4	66.7	65.40	65.6
15	51.1±7.0	50.7	53.2	51.1	51.9	49.49	51.8

16	73.6±4.3	74.9	71.8	73.6	73.6	78.06	71.8
17	38.1±1.7	36.8	37.1	38.1	38.1	31.18	38.6
18	38.3±5.8	39.1	40.5	38.3	39.8	40.10	37.9
19	43.3±3.8	44.1	40.8	43.3	42.7	42.23	42.7
20	54.0±4.5	54.1	51.5	54.1	56.0	54.19	54.9
21	39.2±5.3	40.0	40.4	39.2	39.2	35.83	40.3
22	56.1±1.6	55.7	58.3	56.2	56.3	48.78	55.7
23	58.3±5.4	57.9	57.1	58.3	58.3	57.69	59.3
24	61.0±4.4	60.9	58.5	61.0	60.9	55.04	63.6
25	71.2±5.0	75.0	73	71.2	71.2	74.47	71.2
26	92.4±6.4	89.9	89.1	92.4	93.0	91.90	92.3
27	36.2±2.0	37.7	37.6	36.1	36.2	30.91	37.1
28	44.5±5.2	42.4	42.4	44.5	44.5	42.77	45.0
		CV(RMSE): 6.5%	CV(RMSE): 5.0%	CV(RMSE): 0%	CV(RMSE): 1.3%	CV(RMSE): 6.4%	CV(RMSE): 2.9%
	Avg. of TS&BI	5.7%	5.9%	0%	1.3%	4.4%	2.6%
				AIC: 289	AIC: 227	AIC: 266	AIC: 187
	Avg. of TS&BI			259	194	220	151

RSM: Response Surface Methodology

SVM: Support Vector Machine

BPNN: Backpropagation Neural Network

GenBPNN: Backpropagation Neural Network based on Genetic Algorithm

EvoBPNN: Backpropagation Neural Network based on Evolutionary Algorithm

ELM: Extreme Learning Machine

CV(RMSE): Coefficient of Variance of the Root Mean Squared Error

AIC: Akaike Information Criterion

TS: tensile strength

BI: brittleness index

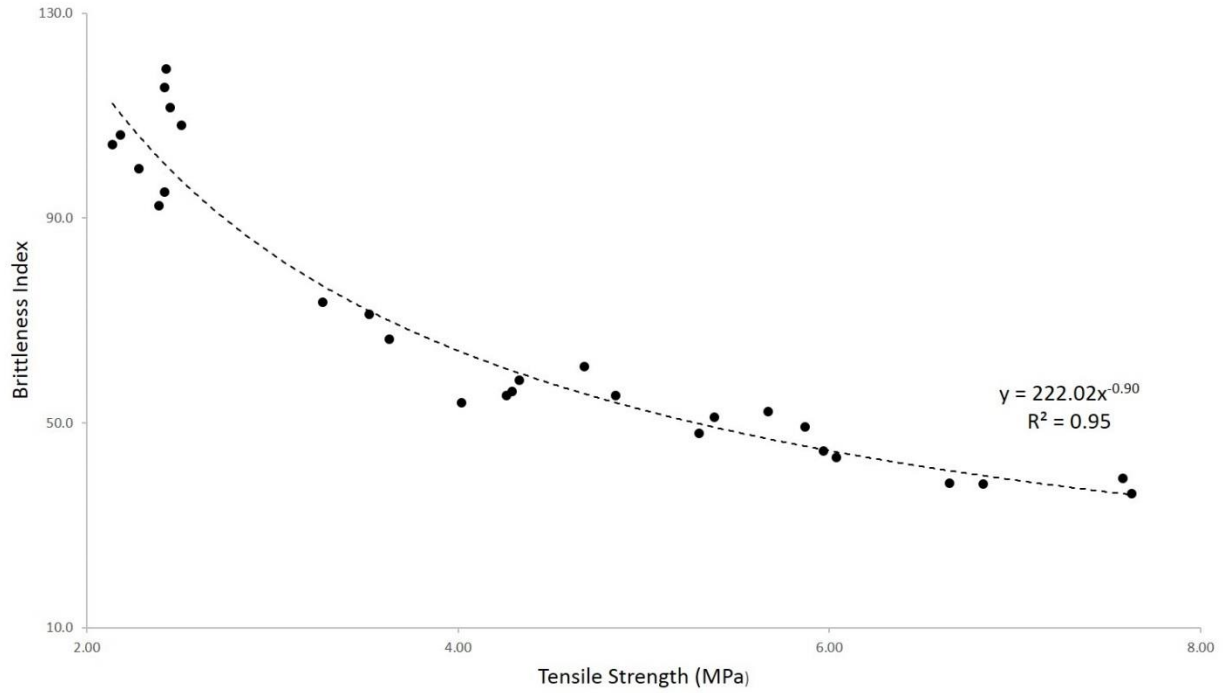

Fig.6. Relationship between brittleness index and tensile strength for the data obtained from DoE studies.

The SEM pictures enabled us to observe the particle surfaces with and without shell. Figure 7(a) and (d) displayed the morphologies of lactose and Calcium Phosphate particles. As seen in Figure 7(b) and (e), for the ratio of Core to Shell at 20/0.1 (% wt/wt), Core/Shell particles exhibited similar morphologies as raw core particles, which meant the core particles were not covered by the shell materials. Conversely, when more shell materials were added in the formula such as the ratio of Core to Shell at 10/2 (% wt/wt), it could be observed from Figure 7(c) and (f) that core powders were mostly or completely coated with one or multiple layers of shell materials and overall the particles' shape was changed.

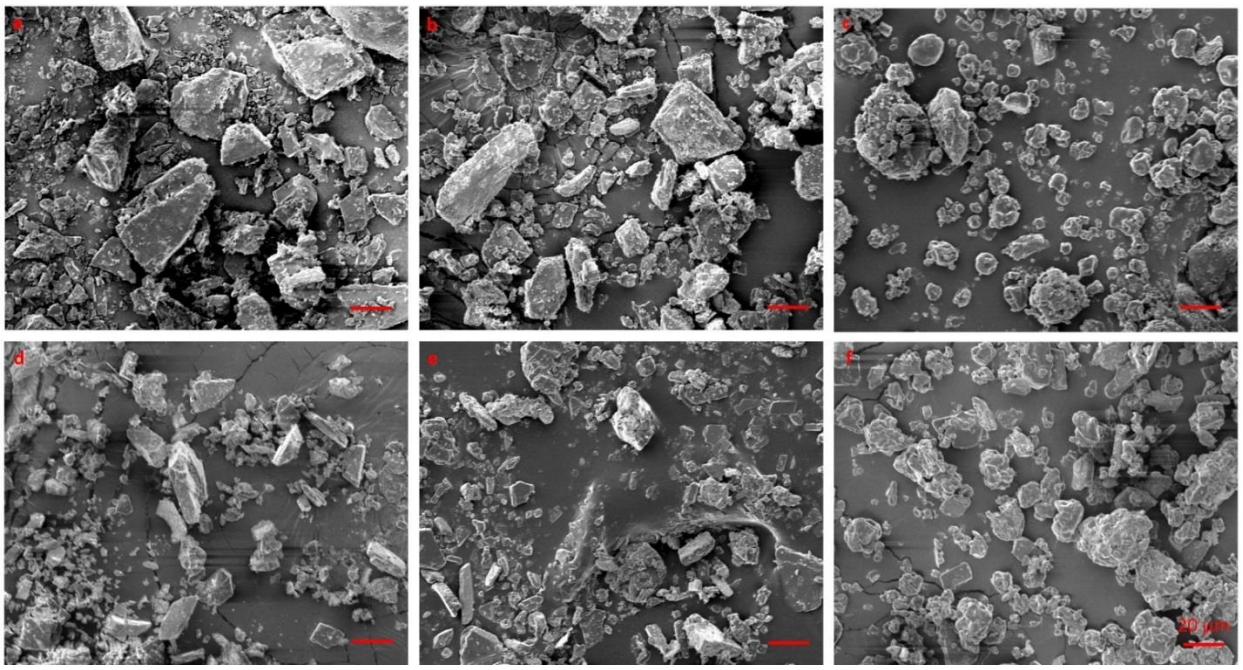

Fig.7. SEM imaginings of samples: (a) Lactose 313; (b) Core/Shell Lactose/VA64 (20/0.1 %wt/wt); (c) Core/Shell Lactose/VA64 (10/2 %wt/wt); (d) Calcium Phosphate Premium Powder; (e) Core/Shell Calcium Phosphate/PVP (20/0.1 %wt/wt); (f) Core/Shell Calcium Phosphate/PVP (10/2 %wt/wt).

3.3.Predictive Modeling and Learning Algorithms

A total of six learning algorithms were used for modeling, including RSM, SVR, BPNN, GA-BPNN, MEA-BPNN, and ELM. As a statistical model, RSM provided a total of 8 equations (shown in Table 4) to account for different combinations of core/shell materials to predict tensile strength or brittleness index values. All models were trained by the dataset from the DoE study and their corresponding predicted results from these models were presented in Tables 5 and 6. All models showed acceptable fitness between observed and predicted values. According to CV(RMSE) values, RSM and SVM models provided the same predictive power. On the other hand, neural networks provided relatively better predictions as indicated by smaller values of

CV(RMSE). Noteworthy, ELM had a CV(RMSE) value of almost 0%, which was in accordance with its algorithm with learning target zero error based on Theorem 1 and 2 proposed by Huang et al [25]. However, we should be extremely cautious to conclude that this ELM model had the best performance, given the likelihood that the model was overfitted with poor generalization performance. In addition, similar AIC values indicated that all the neural networks had similar relative quality.

Type of Material for Shell	Type of Material for Core	Tensile Strength (MPa) (1^{st} equation) Brittleness Index (2^{nd} equation)
S1	C1	$Y = 0.4888 + 5.3651X_1 + 0.2421X_2 - 0.0307X_1X_2 - 1.1882X_1^2 - 0.0086X_2^2$
		$Y = 115.47 - 99.12X_1 + 0.94X_2 + 0.38X_1X_2 + 26.69X_1^2 - 0.04X_2^2$
S1	C2	$Y = 5.4709 + 4.4444X_1 - 0.4329X_2 - 0.0996X_1X_2 - 0.6991X_1^2 + 0.0136X_2^2$
		$Y = 86.57 - 84.11X_1 + 1.79X_2 + 0.90X_1X_2 + 21.91X_1^2 - 0.06X_2^2$
S2	C1	$Y = -8.9535 + 8.1591X_1 + 1.2475X_2 - 0.1679X_1X_2 - 1.7225X_1^2 - 0.0351X_2^2$
		$Y = 219.29 - 143.47X_1 - 9.78X_2 + 1.90X_1X_2 + 36.51X_1^2 + 0.26X_2^2$
S2	C2	$Y = -0.6993 + 1.4114X_1 + 0.4189X_2 + 0.0017X_1X_2 - 0.2226X_1^2 - 0.0144X_2^2$
		$Y = 151.85 - 61.24X_1 - 6.01X_2 + 0.71X_1X_2 + 12.76X_1^2 + 0.20X_2^2$

The trained models were validated using an external dataset. Figures 8 and 9 showed that all points deviated only slightly from dotted line, indicating that the developed models had good generalization performance and validated that the external dataset was well predicted. The tensile strength was better predicted than brittleness index, which may be due to less variation in the tensile strength dataset compared to the brittleness index dataset. Neural networks provided better predictability than RSM and SVR, i.e., for tensile strength, the lowest CV(RMSE): 2.4% from GA-BPNN vs. the highest CV(RMSE): 16.9% from RSM; for brittleness index, the lowest CV(RMSE): 10.4% from MEA-BPNN vs. the highest CV(RMSE): 17.6% from RSM. Also, in principle, neural network algorithms would be expected to continue learning and improving as the dataset becomes larger.

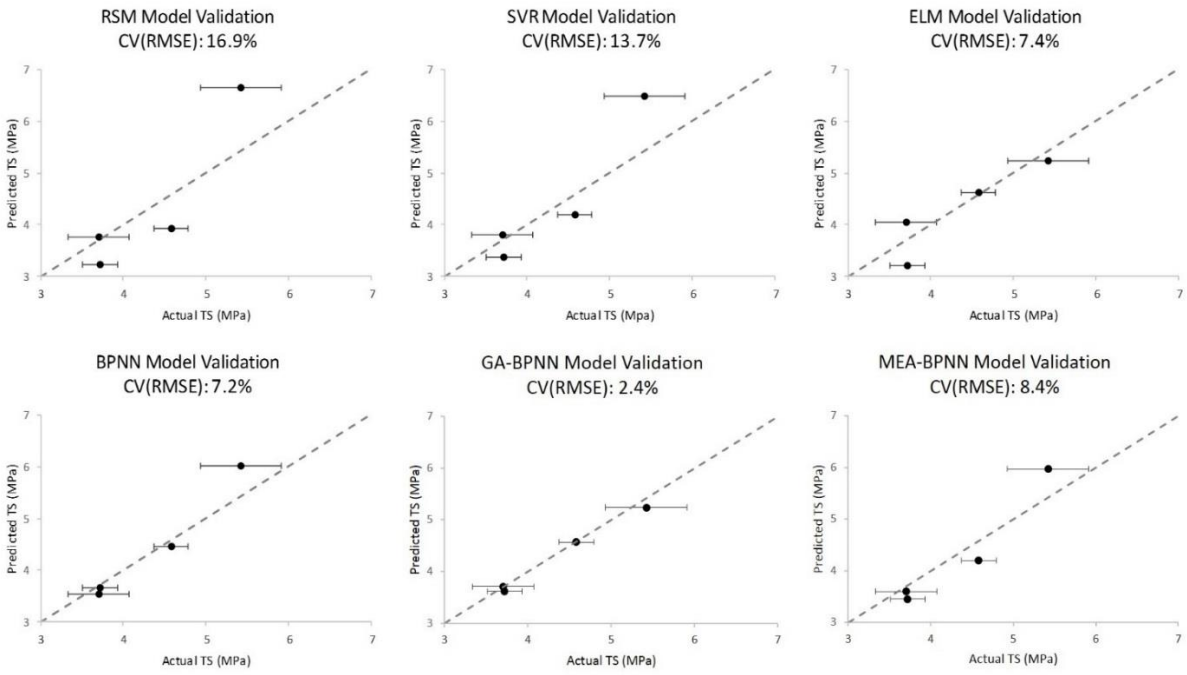

Fig.8. Relationship between observed tensile strength data and predicted tensile strength data from machine learning algorithms as following: RSM, SVR, BPNN, GA-BPNN, MEA-BPNN, and ELM.

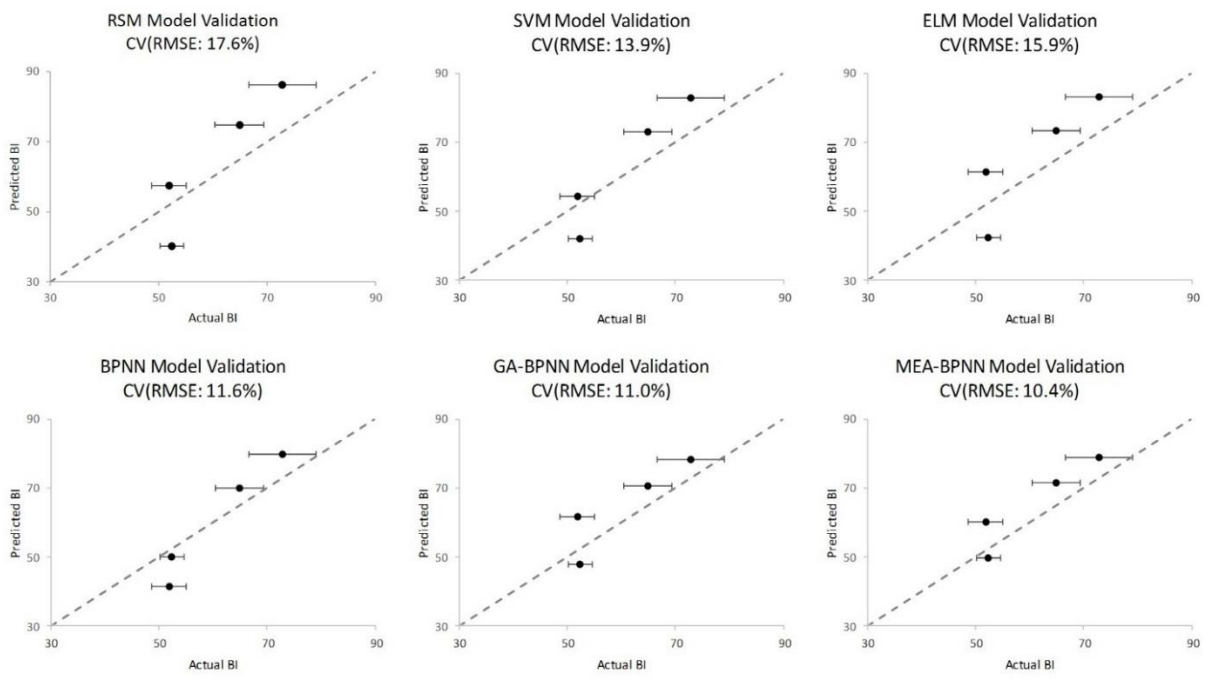

Fig.9. Relationship between observed brittleness index data and predicted brittleness index data from machine learning algorithms as following: RSM, SVR, BPNN, GA-BPNN, MEA-BPNN, and ELM.

Three-dimensional surface plotting showed the relationship of tensile strength or brittleness index and concentrations of core materials and shell materials. To further alleviate any potential outliers created by one model, the mean of the predicted values from BPNN, GA-BPNN, MEA-BPNN, and ELM were calculated and plotted in Figures 10 and 11. Both core material concentration and shell material concentration showed nonlinear effects on tensile strength or brittleness index. Moreover, it was seen that interactions existed between core materials and shell materials and the interaction patterns were different depending on core/shell material types. Overall, PVP served as a better shell material compared to VA64 in terms of improving tensile strength and decreasing brittleness. Also, the compactability of Calcium

Phosphate had more potential to be improved compared to lactose using core/shell technique. Importantly, these plots helped to understand formulation parameters and provide a cost- and time-effective methodology for formulation development and target profile control.

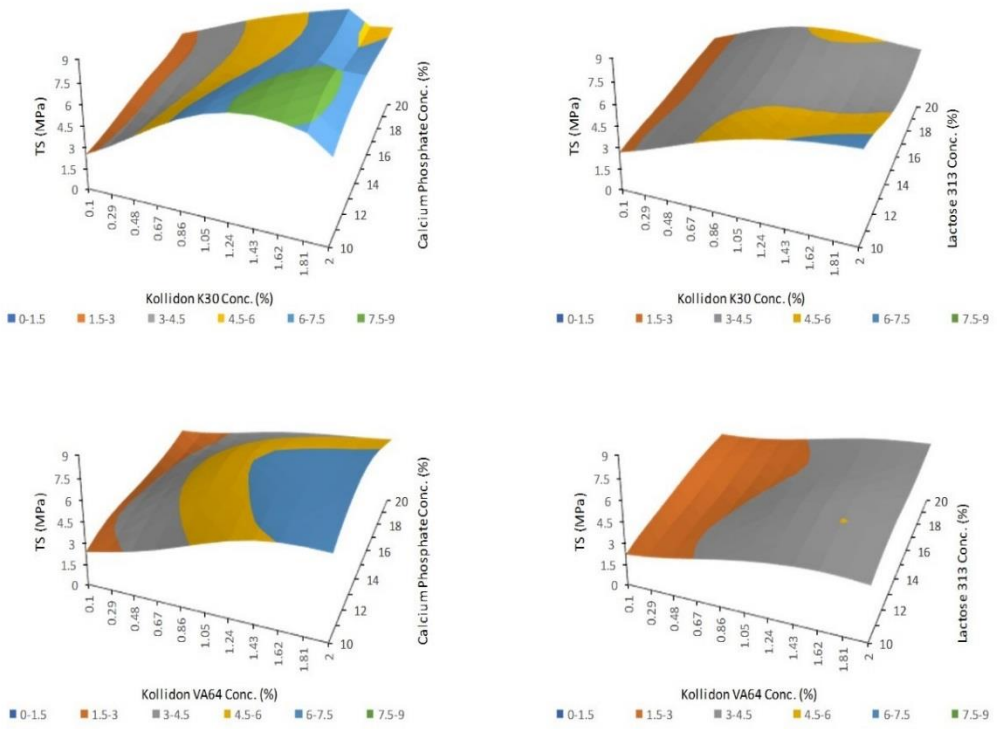

Fig.10. Response surface plot which correlates tensile strength to core/shell material concentrations with different combinations as following: PVP & Calcium Phosphate, PVP & Lactose, VA64 & Calcium Phosphate, VA64 & Lactose.

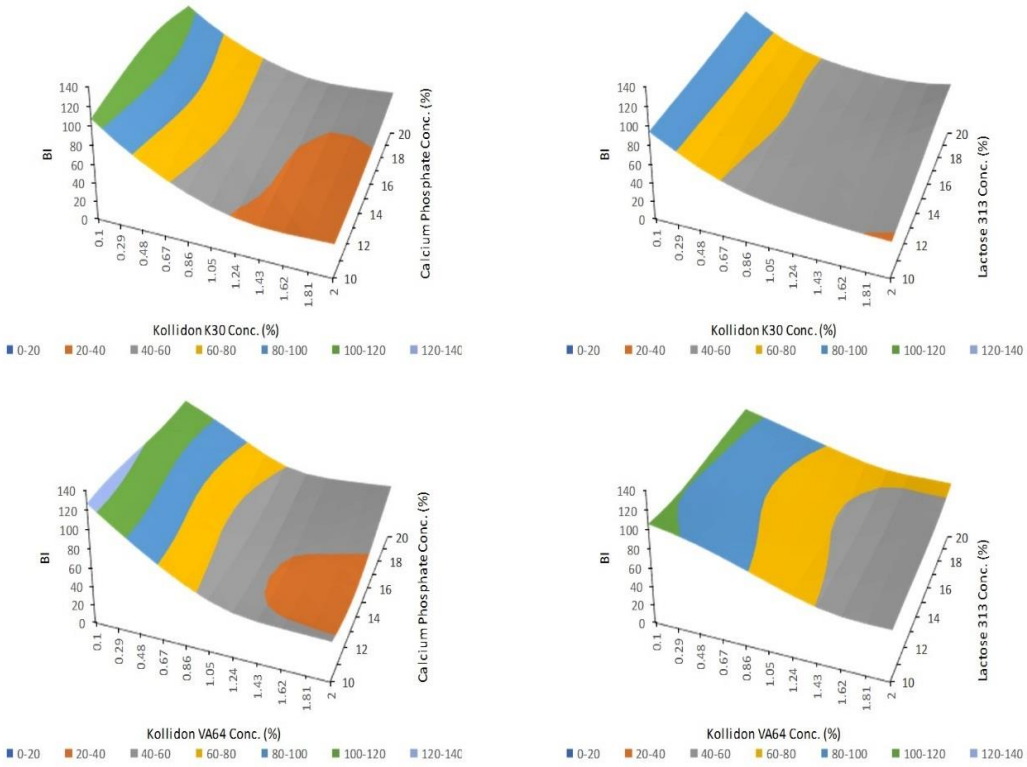

Fig.11. Response surface plot which relates brittleness index to core/shell material concentrations with different combinations as following: PVP & Calcium Phosphate, PVP & Lactose, VA64 & Calcium Phosphate, VA64 & Lactose.

Chapter 4: Conclusion

This study demonstrated that core/shell technique profoundly improved the compactability of poorly compressible powders. This technique and its critical factors were systemically evaluated by a DoE study and the complex relationship between raw material parameters and product properties were modeled using six machine learning algorithms, including RSM, SVR, BPNN, GA-BPNN, MEA-BPNN, and ELM. Overall, these models provided acceptable predictability and capability of generalization and particularly four ANN models were more capable than RSM and SVR. This study provided some insights about how to apply machine learning in pre-formulation and formulation development.

References

[1] C. Sun, D.J.W. Grant, Influence of Crystal Structure on the Tableting Properties of Sulfamerazine Polymorphs, Pharm Res, 18 (2001) 274-280.

[2] C.C. Sun, H. Hou, Improving Mechanical Properties of Caffeine and Methyl Gallate Crystals by Cocrystallization, Crys Growth Des, 8 (2008) 1575-1579.

[3] B.C. Hancock, G.T. Carlson, D.D. Ladipo, B.A. Langdon, M.P. Mullarney, Comparison of the mechanical properties of the crystalline and amorphous forms of a drug substance, Int. J. Pharm., 241 (2002) 73-85.

[4] L. Malaj, R. Censi, Z. Gashi, P. Di Martino, Compression behaviour of anhydrous and hydrate forms of sodium naproxen, Int. J. Pharm., 390 (2010) 142-149.

[5] L. Shi, C.C. Sun, Transforming powder mechanical properties by core/shell structure: Compressible sand, J. Pharm. Sci., 99 (2010) 4458-4462.

[6] L. Shi, C.C. Sun, Overcoming Poor Tabletability of Pharmaceutical Crystals by Surface Modification, Pharm Res, 28 (2011) 3248-3255.

[7] F. Osei-Yeboah, C.C. Sun, Tabletability Modulation Through Surface Engineering, J. Pharm. Sci., 104 (2015) 2645-2648

[8] F. Osei-Yeboah, Y. Lan, C.C. Sun, A top coating strategy with highly bonding polymers to enable direct tableting of multiple unit pellet system (MUPS), Powder Tech, 305 (2017) 591-596.

[9] H. Olsson, C. Nyström, Assessing Tablet Bond Types from Structural Features that Affect Tablet Tensile Strength, Pharm Res, 18 (2001) 203-210.

[10] U. S. Food and Drug Administration. Guidance for Industry: Q8(2) Pharmaceutical Development. 2009

[11] R. Burbidge, M. Trotter, B. Buxton, S. Holden, Drug design by machine learning: support vector machines for pharmaceutical data analysis, Comput. Chem., 26 (2001) 5-14.

[12] E. Byvatov, U. Fechner, J. Sadowski, G. Schneider, Comparison of Support Vector Machine and Artificial Neural Network Systems for Drug/Nondrug Classification, J. Chem. Inf. Comput. Sci., 43 (2003) 1882-1889.

[13] R.N. Jorissen, M.K. Gilson, Virtual Screening of Molecular Databases Using a Support Vector Machine, J. Chem. Inf. Model., 45 (2005) 549-561.

[14] D. Silver, A. Huang, C.J. Maddison, A. Guez, L. Sifre, G. van den Driessche, J. Schrittwieser, I. Antonoglou, V. Panneershelvam, M. Lanctot, S. Dieleman, D. Grewe, J. Nham, N. Kalchbrenner, I. Sutskever, T. Lillicrap, M. Leach, K. Kavukcuoglu, T. Graepel, D. Hassabis, Mastering the game of Go with deep neural networks and tree search, Nature, 529 (2016) 484.

[15] S. Agatonovic-Kustrin, R. Beresford, Basic concepts of artificial neural network (ANN) modeling and its application in pharmaceutical research, J. Pharm. Biomed. Anal., 22 (2000) 717-727.

[16] S.S. Behzadi, C. Prakasvudhisarn, J. Klocker, P. Wolschann, H. Viernstein, Comparison between two types of Artificial Neural Networks used for validation of pharmaceutical processes, Powder Tech, 195 (2009) 150-157.

[17] H. Ichikawa, Hierarchy neural networks as applied to pharmaceutical problems, Adv. Drug. Deliv. Rev., 55 (2003) 1119-1147.

[18] Y. Sun, Y. Peng, Y. Chen, A.J. Shukla, Application of artificial neural networks in the design of controlled release drug delivery systems, Adv. Drug. Deliv. Rev., 55 (2003) 1201-1215.

[19] K. Takayama, M. Fujikawa, Y. Obata, M. Morishita, Neural network based optimization of drug formulations, Adv. Drug. Deliv. Rev., 55 (2003) 1217-1231.

[20] P. Narayan, B.C. Hancock, The influence of particle size on the surface roughness of pharmaceutical excipient compacts, Mat. Sci. Eng. A., 407 (2005) 226-233.

[21] X. Gong, C.C. Sun, A new tablet brittleness index, Euro. J. Pharm. Biopharm., 93 (2015) 260-266.

[22] A.J. Smola, B. Schölkopf, A tutorial on support vector regression, Stat. Comp., 14 (2004) 199-222.

[23] J. Jie, J. Zeng, C. Han, An extended mind evolutionary computation model for optimizations, Appl. Math. Comp., 185 (2007) 1038-1049.

[24] H. Liu, H. Tian, X. Liang, Y. Li, New wind speed forecasting approaches using fast ensemble empirical model decomposition, genetic algorithm, Mind Evolutionary Algorithm and Artificial Neural Networks, Renew Energy, 83 (2015) 1066-1075.

[25] G.-B. Huang, Q.-Y. Zhu, C.-K. Siew, Extreme learning machine: Theory and applications, Neurocomputing, 70 (2006) 489-501.

[26] C.C. Sun, A classification system for tableting behaviors of binary powder mixtures, Asia. J. Pharm. Sci., 11 (2016) 486-491.

[27] X. Gong, S.-Y. Chang, F. Osei-Yeboah, S. Paul, S.R. Perumalla, L. Shi, W.-J. Sun, Q. Zhou, C.C. Sun, Dependence of tablet brittleness on tensile strength and porosity, Int. J. Pharm., 493 (2015) 208-213.

[28] K.-i. Sugimori, Y. Kawashima, H. Takeuchi, T. Hino, T. Niwa, S. Ohno, S. Mori, Effects of Granulation Method and Drug Dissolved in Binder Solution on Compressibility of Granules, Chem. Pharm. Bull., 38 (1990) 188-192.

[29] Q. Zhou, L. Qu, I. Larson, P.J. Stewart, D.A.V. Morton, Effect of mechanical dry particle coating on the improvement of powder flowability for lactose monohydrate: A model cohesive pharmaceutical powder, Powder Technology, 207 (2011) 414-421.

[30] Q. Zhou, L. Shi, W. Marinaro, Q. Lu, C.C. Sun, Improving manufacturability of an ibuprofen powder blend by surface coating with silica nanoparticles, Powder Technology, 249 (2013) 290-296.